International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

Machine Learning predicting Indian GDP using Machine learning

J. SASIREKHA, Department Of IT, SICET, Hyderabad
MANDLAM SAIABHINAYA REDDY, ALISHETTY NITHIN, KANURU RATHNA PRAKASHRAJ, VUTHURI VARSHA
UG Student, Department Of IT, SICET, Hyderabad

Abstract

The significance of GDP may be observed in the fact that it gives information on the size and performance of an economy. The rate of growth of real GDP is widely used as an indicator of the economy is performing well. This project aims is to predict the world's GDP by using machine learning and also calculate their year to year growth .Every year to year GDP is fluctuating it is important to know the current and importance of the factors which is affected for the country's GDP. By showing this thing in a graph it is easy to understand. This research based project also calculating the importance of features affected for the calculation of the GDP. It is very useful for the viewer to view the GDP growth of the different countries and also they can see the best and worst predict performance.

Keywords– GDP, Factors, Predictor.

INTRODUCTION

Gross Domestic Product is the total monetary value of the all the finished goods and services produced within a country in a specific time period.Countries with higher GDPs will have a greater emphasis on the labour and goods produced inside their borders, and will, on average, have a higher standard of life. The gross domestic product is usually calculated on an annual basis, although it may also be calculated quarterly. Typically the Gross Domestic Product or service (GDP) is vital since it provides information on the type and performance of an economy.

Typically the project "Global GROSS DOMESTIC PRODUCT Prediction" is structured on research. Typically the goal of this project is to forecast the GROSS DOMESTIC PRODUCT per capita of various countries, as well as being the relevance of the factors that influence GROSS DOMESTIC PRODUCT calculation. We can achieve the best and worst prediction performance by evaluating four different learning regressions (Linear Regression, SVM, Random Forest, and Gradient Boosting) in this study. Our research's main objective is to use Machine Learning and Python to anticipate GDP growth.

In this work we have used the library numPy for working with arrays, pandas used to perform data analysis and manipulation, and the matplotlib library for plot the points. We used four different learningregressions: linear regression, support vector machine, random forest, and gradient boosting.

I. LITERATURE SURVEY

"World's GDP prediction" is a researched based project. It is important to everyone to know about their country's GDP. By using some algorithm here I calculated importance of features and how much they affect in calculation of GDP and in the next few what will be our and other countries economic growth. For doing these I referred some papers to gain more knowledge about this.

In this paper author explain about the GDP from the basic points and their types, ways of calculating GDP, advantages and disadvantages.[1].

This research describes regression and provides comprehensive information on linear regression, simple and multiple linear regression, implementing linear regression in Python, python libraries for linear regression, and simple and multiple linear regression with scikit-learn.[2].

In this Paper they gone through about support vector machine, what is SVM, how that algorithm works and some information about support vector regression, implementing SVR in python which includes importing libraries, reading the dataset and feature scaling.[3].

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

In this study author explains about the Random forest algorithm, how to import libraries and dataset, how that dataset splits, how that predict the results.[4].

The facts regarding gradient boosting, the genesis of boosting from learning theory, and adaboost will be provided in this work. The loss function, weak learners, and the additive model all play a role in how gradient boosting works. How to use multiple regularisation strategies to increase performance over the base algorithm.[5].

This research looks at how machine learning may be used to forecast economic variables. Although AI has been integrated into economics, machine learning projections have yet to be completely tested. In the instance of current cases from G7 countries, a comparison of employing AR models and machine learning to anticipate GDP and consumer price is undertaken. [6]

In this paper it will provide the information about India's economy, what are all the factors affecting, how we can calculate the GDP, how we can predict the data.[11]

This paper give the information about the projected real GDP growth, relative income analysis, scenario analysis, economic growth in2050.[12].

II. PROPOSED SYSTEM

In this researched based GDP predictor we have collect the one dataset which has almost all the country's factors value for the calculation of the GDP. In that we are selecting some of the country's name and showing that country's GDP results in graphs and also showing the importance of factors, And using 4 linear regressors we are calculating Mean Absolute Error, Root mean squared error (RMSE) and R-squared Score (R2_Score) and shows that result in graph.

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

This paper joins a developing writing that assesses the general achievement of ML models in estimating throughout the more conventional time-series procedures. In this paper, we examine the presentation of various ML calculations in acquiring precise now projects of genuine total national output (GDP) development for New Zealand. We utilize various vintages of authentic GDP information and numerous vintages of an enormous highlights set. [7]

Our goal is to use multiple sophisticated machine learning techniques to develop and evaluate a forecasting model for GDP changes. We investigate if machine learning algorithms can increase forecasting accuracy and attempt to identify the factors that drive economic recovery and predict recession.[8]
In this paper, we explore the exhibition of various ML calculations in getting exact now projects of the current quarter genuine (GDP) development for New Zealand. We utilize various vintages of recorded GDP information and different vintages of an enormous highlights set - involving roughly 550 homegrown and worldwide factors - to assess the continuous exhibition of these calculations over the 2009-2018 period.[9].

The primary goal of this article is to examine the trajectory of GDP to determine if there is any change in the trend between the years 1981-82 to 1990-91 and 1991-92 to 2001-2002, i.e. two decades before and after the liberalisation programme began. 2. Investigate the nature of the difference, if one exists. 3. Determine the reasons for the difference. 4. Create a GDP forecasting model.[10].

WORKING

In this work we collect one data set which include all the factors value which is used for GDP calculation. In this work we have also calculated the importance factors and shows it in a graph like which factors affect the most.
In this first we have to extract the dataset which stored as a CSV file then we have to check the factors value if it contains any zero value or factors contains any other data types. Then using some machine learning algorithm current and prediction of GDP will show in a graph.

III. IMPLEMENTATION

- This section contains all of the information on the technology employed and the project's control flow. This project is built with Python, which is developed in Python 3.8.7 and is simple for developers to comprehend and implement.
- Install Python or Open Colaboratory
- Create a project by giving the name of the project.

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

clear, logical code for both small and big projects. Python is garbage-collected and dynamically typed. It supports a variety of programming diagrams, including procedural, object-oriented, and functional

programming. Python is sometimes referred to as a "batteries included" language because of its large standard library.

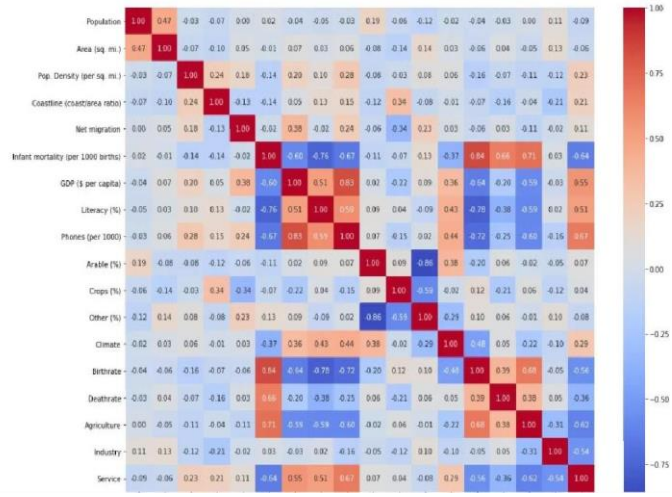

Figure 1: Corelationheatmap

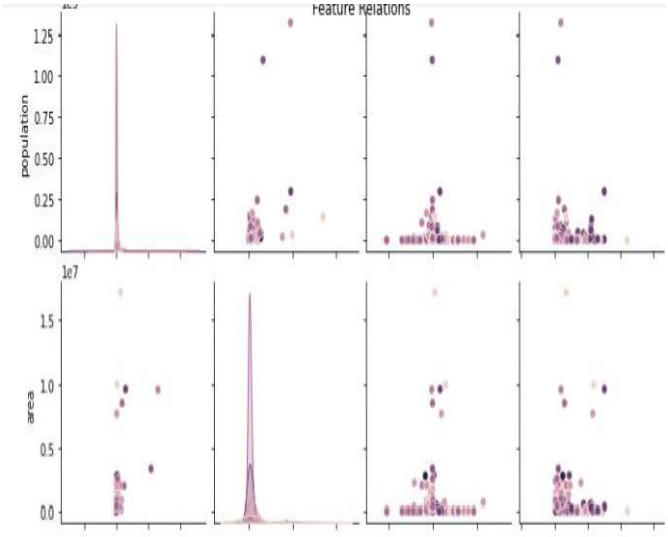

Figure 2: Feature relation

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

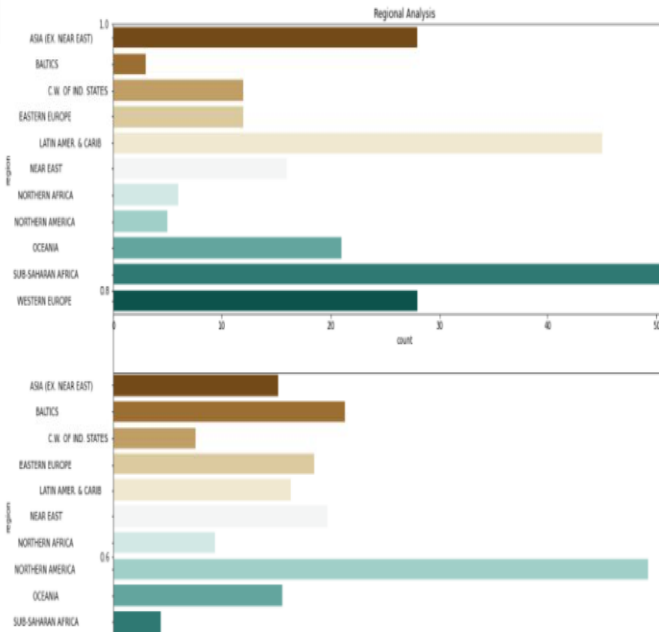

Figure 4: Regional analysis

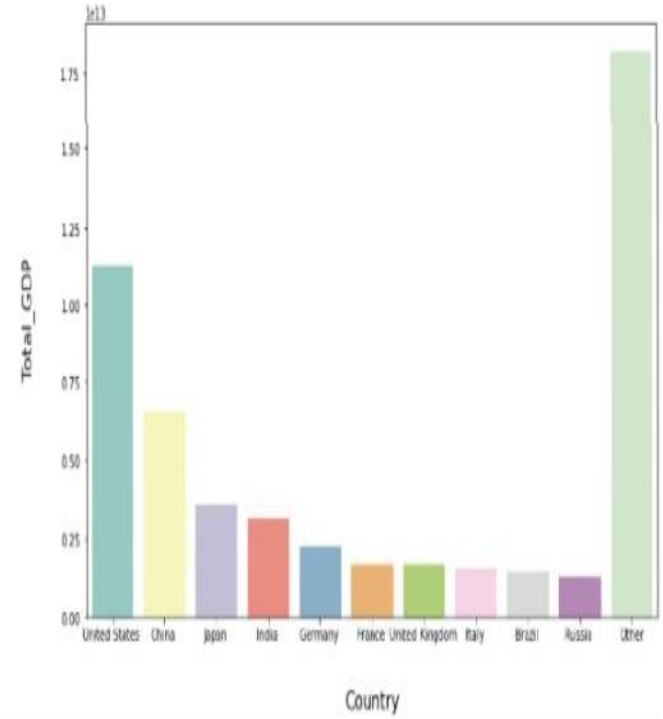

Figure 5: Different country's GDP growth

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

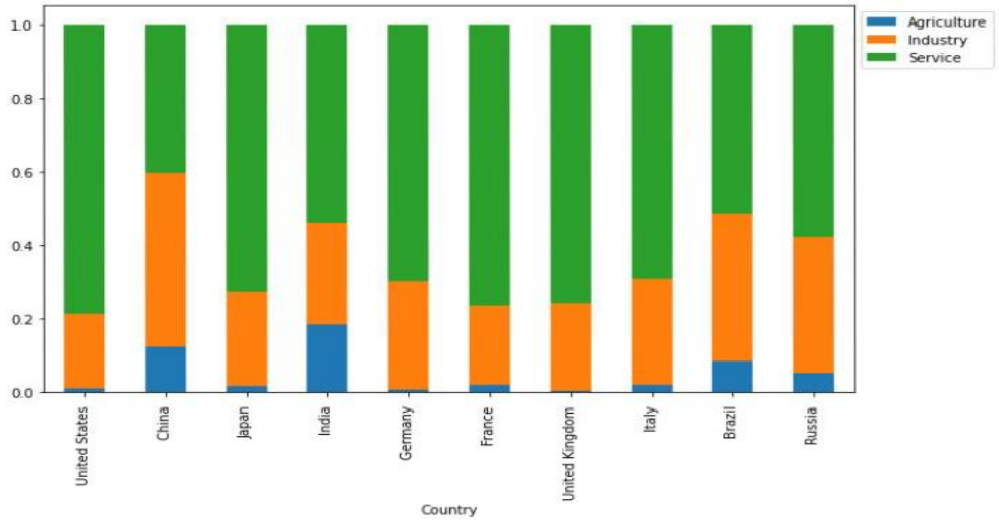

Figure 6: Importance of Feature

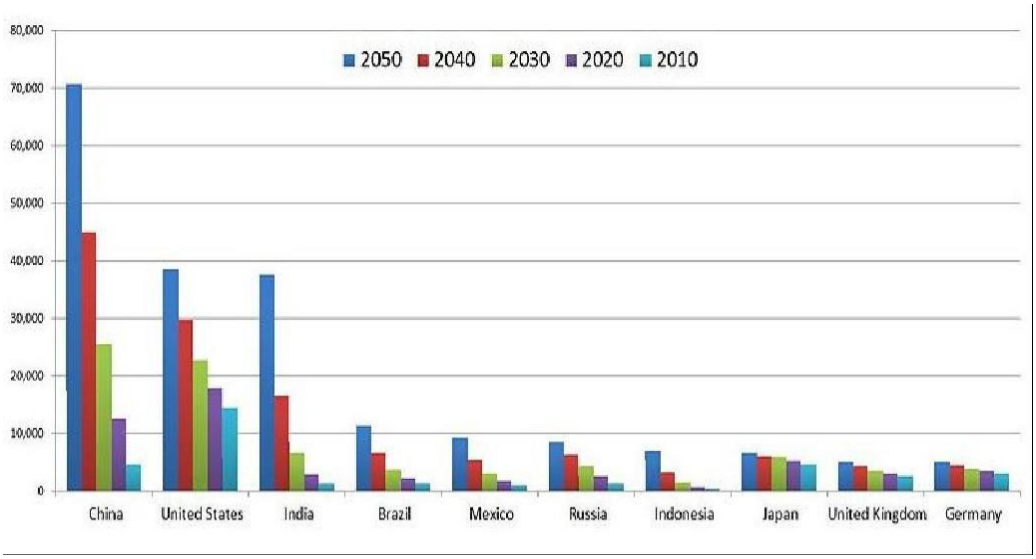

Figure 7: Prediction of GDP

IV. RESULT ANALYSIS

Although we can see from our EDA that most characteristics do not have a linear connection with our labels (gdp per capita), we will attempt linear regression and utilise the outcome as a guide (other methods should have better results).

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

It is clear that Gradient Boosting gave us good performance even before optimization. Its performance on our dataset is very close to that of Random Forest.The best prediction performance was achieved using **Random Forest** regression, using all features in the dataset, and resulted in the ollowing metrics:

- Mean Absolute Error **(MAE)**: 2142.13
- Root mean squared error **(RMSE)**: 3097.19
- R-squared Score **(R2_Score)**: 0.8839

Taking into account that the gdp_per_capita values in the dataset ranges from 500 to 55100 USD.

V. CONCLUSION

In this project we have developed a predictor which is easy to understand our country or different country's GDP. It is important to know and to have the some basic information about the GDP. In this project viewer can easily understand the importance of features and how much they affected on the GDP, also the best and worst predictor performance using linear regressors.

.

VI. REFRENCES

[1] Fernando J. Gross Domestic Product (GDP). Investopedia. 2021.

[2] Laine M. Introduction to dynamic linear models for time series analysis. InGeodetic Time Series Analysis in Earth Sciences 2020 (pp. 139-156). Springer, Cham.

[3] Lalitha VL, Raju SH, Sonti VK, Mohan VM. Customized Smart Object Detection: Statistics of detected objects using IoT. In2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) 2021 Mar 25 (pp. 1397-1405). IEEE.

[4] Qin T. Machine Learning Basics. InDual Learning 2020 (pp. 11-23). Springer, Singapore.

[5] Brownlee J. A Gentle Introduction to the Gradient Boosting Algorithm for Machine Learning. 2016.

[6] Kurihara Y, Fukushima A. AR Model or Machine Learning for Forecasting GDP and Consumer Price for G7 Countries. Applied Economics and Finance. 2019;6(3):1-6.

[7] Richardson A, van Florenstein Mulder T, Vehbi T.

Nowcasting GDP using machine-learning algorithms: A real-time assessment. International Journal of Forecasting. 2021 Apr 1;37(2):941-8.

[8] Cicceri G, Inserra G, Limosani M. A machine learning approach to forecast economic recessions—an italian case study. Mathematics. 2020 Feb;8(2):241.

[9] Richardson A, Mulder T. Nowcasting New Zealand GDP using machine learning algorithms.

[10] Daga UR, Das R, Maheshwari B. Estimation, Analysis and Projection of India's GDP.

[11] Chen DH, Dahlman CJ. The knowledge economy, the KAM methodology and World Bank operations. World Bank Institute Working Paper. 2005 Oct 19(37256).

[12] Hawksworth J, Cookson G. The world in 2050. How big will the major emerging market economies get and how can the OECD compete. 2006 Sep.